

## Guidelines

# REporting recommendations for tumour MARKer prognostic studies (REMARK)

**LM McShane<sup>\*,1</sup>, DG Altman<sup>2</sup>, W Sauerbrei<sup>3</sup>, SE Taube<sup>1</sup>, M Gion<sup>4</sup> and GM Clark<sup>5</sup> for the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics<sup>6</sup>**

<sup>1</sup>US National Cancer Institute, Bethesda, MD 20892, USA; <sup>2</sup>Cancer Research UK Medical Statistics Group, Centre for Statistics in Medicine, Wolfson College, Oxford OX2 6UD, UK; <sup>3</sup>Institut fuer Medizinische Biometrie und Medizinische Informatik, Universitaetsklinikum Freiburg, 79104 Freiburg, Germany; <sup>4</sup>Centro Regionale Indicatori Biochimici di Tumore, Ospedale Civile, 30122 Venezia, Italy; <sup>5</sup>OSI Pharmaceuticals, Inc., Boulder, CO 80301, USA

Despite years of research and hundreds of reports on tumour markers in oncology, the number of markers that have emerged as clinically useful is pitifully small. Often initially reported studies of a marker show great promise, but subsequent studies on the same or related markers yield inconsistent conclusions or stand in direct contradiction to the promising results. It is imperative that we attempt to understand the reasons that multiple studies of the same marker lead to differing conclusions. A variety of methodological problems have been cited to explain these discrepancies. Unfortunately, many tumour marker studies have not been reported in a rigorous fashion, and published articles often lack sufficient information to allow adequate assessment of the quality of the study or the generalisability of the study results. The development of guidelines for the reporting of tumour marker studies was a major recommendation of the US National Cancer Institute and the European Organisation for Research and Treatment of Cancer (NCI-EORTC) First International Meeting on Cancer Diagnostics in 2000. Similar to the successful CONSORT initiative for randomised trials and the STARD statement for diagnostic studies, we suggest guidelines to provide relevant information about the study design, preplanned hypotheses, patient and specimen characteristics, assay methods, and statistical analysis methods. In addition, the guidelines suggest helpful presentations of data and important elements to include in discussions. The goal of these guidelines is to encourage transparent and complete reporting so that the relevant information will be available to others to help them to judge the usefulness of the data and understand the context in which the conclusions apply.

*British Journal of Cancer* (2005) **93**, 387–391. doi:10.1038/sj.bjc.6602678 www.bjcancer.com

Published online 2 August 2005

© 2005 Cancer Research UK

**Keywords:** tumour marker; guidelines; REMARK; NCI; EORTC; prognostic

\*Correspondence: Dr LM McShane, National Cancer Institute, Biometric Research Branch, DCTD, Room 8126, Executive Plaza North, MSC 7434, 6130 Executive Boulevard, Bethesda, MD 20892-7434, USA; E-mail: Lm5h@nih.gov

In order to encourage dissemination of the guidelines set out in this paper, it is freely accessible on the BJC website (www.bjcancer.com), and will also be published simultaneously in the *European Journal of Cancer*, *Journal of Clinical Oncology*, *Journal of the National Cancer Institute*, and *Nature Clinical Practice Oncology*.

<sup>6</sup>Members of the Statistics Subcommittee of the NCI/EORTC Working Group on Cancer Diagnostics: Douglas G Altman DSc (Co-chair), Cancer Research UK Medical Statistics Group, Centre for Statistics in Medicine, Wolfson College, Oxford OX2 6UD, UK; Lisa M McShane, PhD (Co-chair), US National Cancer Institute, Bethesda, MD 20892, USA; Gary M Clark, PhD, OSI Pharmaceuticals, Inc., Boulder, CO 80301, USA; Jose Costa, MD, Yale Cancer Center, New Haven, CT 06510-3202, USA; Angelo Di Leo, MD, PhD, Department of Oncology, Hospital of Prato, 59100 Prato, Italy; Massimo Gion, MD, Centro Regionale Indicatori Biochimici di Tumore, Ospedale Civile, 30122 Venezia, Italy; Robert J Mayer, MD, Dana-Farber Cancer Institute, Boston, MA 02115, USA; Willi Sauerbrei, PhD, Institut fuer Medizinische Biometrie und Medizinische Informatik, Universitaetsklinikum Freiburg, 79104 Freiburg, Germany; Sheila E Taube, Ph.D., US National Cancer Institute, Bethesda, MD 20892, USA

Received 5 January 2005; revised 18 May 2005; accepted 25 May 2005; published online 2 August 2005

Despite years of research and hundreds of reports on tumour markers in oncology, the number of markers that have emerged as clinically useful is pitifully small (Hayes *et al*, 1996; Bast *et al*, 2001; Schilsky and Taube, 2002). Often initially reported studies of a marker show great promise, but subsequent studies on the same or related markers yield inconsistent conclusions or stand in direct contradiction to the promising results. It is imperative that we attempt to understand the reasons that multiple studies of the same marker lead to differing conclusions. A variety of problems have been cited to explain these discrepancies, such as general methodological differences, poor study design, assays that are not standardised or lack reproducibility, and inappropriate or misleading statistical analyses often based on sample sizes too small to draw meaningful conclusions (McGuire, 1991; Fielding *et al*, 1992; Burke and Henson, 1993; Concato *et al*, 1993; Gasparini *et al*, 1993; Simon and Altman, 1994; Gasparini, 1998; Hall and Going, 1999). For example, in retrospective studies, patient populations are often biased towards patients with available tumour specimens. Specimen availability may be related to tumour size and patient outcome (Hoppin *et al*, 2002), and the quantity, quality, and preservation method of the specimen may affect feasibility of conducting certain assays. There can also be biases or large variability inherent in the assay results, depending on the particular assay methods used (Thor *et al*, 1999; Gancberg *et al*, 2000; McShane *et al*, 2000; Paik *et al*, 2002; Roche *et al*, 2002). Statistical problems are commonplace. These problems include

underpowered studies or overly optimistic reporting of effect sizes and significance levels due to multiple testing, subset analyses, and cutpoint optimisation (Altman *et al*, 1995).

Unfortunately, many tumour marker studies have not been reported in a rigorous fashion, and published articles often lack sufficient information to allow adequate assessment of the quality of the study or the generalisability of study results. Such reporting deficiencies are increasingly being highlighted by systematic reviews of the published literature on particular markers or cancers (Brundage *et al*, 2002; Mirza *et al*, 2002; Riley *et al*, 2003a, b, 2004; Burton and Altman, 2004; Popat *et al*, 2004).

The development of guidelines for the reporting of tumour marker studies was a major recommendation of the US National Cancer Institute and the European Organisation for Research and Treatment of Cancer (NCI-EORTC) First International Meeting on Cancer Diagnostics (From Discovery to Clinical Practice: Diagnostic Innovation, Implementation, and Evaluation) that was convened in Nyborg, Denmark in July 2000. The purpose of the meeting was to discuss issues, accomplishments, and barriers in the field of cancer diagnostics. Poor study design and analysis, assay variability, and inadequate reporting of studies were identified as some of the major barriers to progress in this field. One of the working groups formed at the Nyborg meeting was charged with addressing statistical issues of poor design and analysis, and reporting of tumour marker prognostic studies. The guidelines presented here are the product of that committee. The Program for the Assessment of Clinical Cancer Tests (PACCT) Strategy Group of the US National Cancer Institute has also strongly endorsed this effort (<http://www.cancerdiagnosis.nci.nih.gov/assessment/>).

The reporting guidelines proposed in this paper build upon earlier suggestions (Altman and Lyman, 1998; Gion *et al*, 1999; Altman, 2001a, b; Riley *et al*, 2003a) as well as educational publications (McShane and Simon, 2001; Simon, 2001; Biganzoli *et al*, 2003; Schumacher *et al*, 2005). They recommend elements and formats for presentation with the objectives of facilitating evaluation of the appropriateness and quality of study design, methods, analyses, and improving the ability to compare results across studies. Similar to the successful CONSORT initiative for randomised clinical trials (Moher *et al*, 2001), and the STARD statement for studies of diagnostic test accuracy (Bossuyt *et al*, 2003a), these guidelines suggest relevant information that should be provided about the study design, preplanned hypotheses, patient and specimen characteristics, assay methods, and statistical analysis methods. In addition, the guidelines suggest helpful presentations of data and important elements to include in discussions. To be published separately, in an explanatory document, are specific justifications for the need for each of the elements of the recommendations.

We have developed these reporting guidelines primarily for studies evaluating a single tumour marker of interest, often including adjustment for standard clinical prognostic variables. They are largely relevant for studies exploring more than one marker, but they are not intended to specifically address statistical considerations in development of prognostic models from very large numbers of candidate markers. The reason we chose to emphasise prognostic marker studies is that they represent a large proportion of the tumour marker literature and tend to be particularly fraught with problems because they are often conducted on retrospective collections of specimens, and analyses may contain substantial exploratory components. For purposes of this paper, we define prognostic markers to be markers that have an association with some clinical outcome, typically a time-to-event outcome such as overall survival or recurrence-free survival. (Some individuals adhere to a more strict definition of prognostic marker as applying only to the natural history of patients who received no treatment following local therapy.) Prognostic markers may be considered in the clinical management of a patient. For

example, they may be used as decision aids in determining whether a patient should receive adjuvant chemotherapy or how aggressive that therapy should be. Predictive markers are generally used to make more specific choices between treatment options. Predictive markers are used as indicators of the likely benefit to a specific patient of a specific treatment. For example, a predictive marker might indicate that a patient expressing the marker will benefit more from a new treatment compared to standard treatment, whereas a patient not expressing the marker will derive little or no benefit from the new treatment. Predictive marker studies usually occur later in the marker development process and there are far fewer published examples. Knowledge of specific treatments received and how those treatment decisions were made become even more critical. In our judgment, the issues in reporting predictive marker studies are complex and different enough from those of prognostic marker studies that we are not willing to claim that these guidelines give predictive marker studies adequate coverage, although we believe that most of the guidance is relevant to such studies too.

The goal of these guidelines is to encourage transparent and complete reporting so that the relevant information will be available to others to help them to judge the usefulness of the data and understand the context in which the conclusions apply. These guidelines are not intended to dictate specific designs or analysis strategies. In general, there is more than one acceptable approach to the design or analysis of a particular study, although these guidelines should help to eliminate some clearly unacceptable options as have been discussed in other papers (Concato *et al*, 1993; Altman *et al*, 1994; Altman and Lyman, 1998; Schumacher *et al*, 2005). For example, unacceptable options include reporting statistical significance of a marker's prognostic effect without acknowledging that the significance testing was preceded by extensive manipulations involving derivation of data-dependent cutpoints or variable selection procedures. High-quality reporting of a study cannot transform a poorly designed or analysed study into a good one, but it can help to identify the poor studies and we believe that it is an important first step in improving the overall quality of tumour marker prognostic studies.

## MATERIALS AND METHODS

Initial ideas for key elements to be addressed in the guidelines were assembled from literature citing empirical evidence of inadequate reporting or problematic analysis methods (Hilsenbeck *et al*, 1992; Altman *et al*, 1994, 1995; Simon and Altman, 1994) based on published reviews of tumour marker studies. Ideas were also generated by reviewing similar reporting guidelines that have been produced for other types of medical research studies (CONSORT, QUOROM, MOOSE, STARD) (Moher *et al*, 1999, 2001; Stroup *et al*, 2000; Bossuyt *et al*, 2003a). Three individuals from the working group (LM, DA, GC) wrote a first draft to serve as a starting point for discussion by the full group. Comments on drafts were made by the full group on a conference call and through multiple e-mail exchanges. A very preliminary draft was presented to the PACCT Strategy Group in January 2001. In response to comments, the guidelines were shortened, reformatted, and recirculated to the full committee. They were posted to the PACCT website (<http://www.cancerdiagnosis.nci.nih.gov/assessment/progress/clinical.html>) for public comment and circulated to attendees of the NCI-EORTC Second International Meeting on Cancer Diagnostics (Conference on the Development of New Diagnostic Tools for Cancer) that was held in Washington, DC in June 2002. In February 2003, three committee members (DA, LM, WS) met for 2 days to make further revisions. The version produced in that February meeting was sent to the full committee for final comment. The version presented here incorporates those final comments and was approved by the full committee.

## RESULTS

Table 1 shows the recommendations for reporting studies on tumour markers. Specific items are grouped under headings: Introduction, Materials and Methods, Results, and Discussion, reflecting the relevant sections of a published scientific article. Further details about the recommendations and explanatory material will be provided in a separate article.

As noted in item 12, a diagram may be helpful to indicate numbers of individuals included at different stages of a study. As a minimum, such a diagram could show the number of patients originally in the sample, the number remaining after exclusions, and the numbers incorporated into univariate and multivariable analyses.

## DISCUSSION

The reporting guidelines presented here are the result of a collaborative effort among statisticians, clinicians, and laboratory scientists who are committed to improving and accelerating the process by which tumour markers that provide useful information

for management of cancer patients are adopted into clinical practice. In addition to the authors of this paper, we gratefully acknowledge the contributions of many individuals with whom we have had informal discussions regarding these guidelines and who have been supportive of this effort. All of us participating in the development of these guidelines are actively involved in the design, conduct, and analysis of studies involving tumour markers. Collectively, we serve as editors and reviewers for numerous scientific journals that publish tumour marker studies, we serve on programme committees for international meetings, as decision-makers for funding agencies, participants in national and international committees charged with evaluating and prioritising tumour markers for further study or making recommendations for clinical use, and are actively involved in our own research involving tumour markers. As editors, reviewers, and programme and advisory committee members, we have struggled with having to make decisions when insufficient information is provided about study design or analysis methods. As individual investigators, we have experienced the frustration of trying to interpret often confusing literature to guide our own research programmes.

**Table 1** REporting recommendations for tumour MARKer prognostic studies (REMARK)

### Introduction

1. State the marker examined, the study objectives, and any prespecified hypotheses.

### Materials and Methods

#### Patients

2. Describe the characteristics (e.g. disease stage or comorbidities) of the study patients, including their source and inclusion and exclusion criteria.
3. Describe treatments received and how chosen (e.g. randomised or rule-based).

#### Specimen characteristics

4. Describe type of biological material used (including control samples), and methods of preservation and storage.

#### Assay methods

5. Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study end point.

#### Study design

6. State the method of case selection, including whether prospective or retrospective and whether stratification or matching (e.g. by stage of disease or age) was employed. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.
7. Precisely define all clinical end points examined.
8. List all candidate variables initially examined or considered for inclusion in models.
9. Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.

#### Statistical analysis methods

10. Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.
11. Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.

## Results

### Data

12. Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the numbers of patients and the number of events.
13. Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumour marker, including numbers of missing values.

### Analysis and presentation

14. Show the relation of the marker to standard prognostic variables.
15. Present univariate analyses showing the relation between the marker and outcome, with the estimated effect (e.g. hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analysed. For the effect of a tumour marker on a time-to-event outcome, a Kaplan–Meier plot is recommended.
16. For key multivariable analyses, report estimated effects (e.g. hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.
17. Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their significance.
18. If done, report results of further investigations, such as checking assumptions, sensitivity analyses, internal validation.

### Discussion

19. Interpret the results in the context of the prespecified hypotheses and other relevant studies; include a discussion of limitations of the study.
20. Discuss implications for future research and clinical value.

There are consequences of poor study reporting for the research community as a whole. Poorly designed or inappropriately analysed studies can attract undeserved attention when they produce very dramatic, but unfortunately incorrect results. In contrast, some carefully designed and analysed studies have been overlooked because they produced less dramatic, but perhaps more accurate and realistic results. The poor quality of reporting of prognostic marker studies may have contributed to the relative scarcity of markers whose prognostic influence is well-supported. Thorough reporting is required no matter what methods of design and analysis are used. Thorough reporting does not solve problems of poor design or analysis that are being reported; rather, it just fairly describes what problems may exist and need to be considered in interpretation. It is our hope that these guidelines will be embraced and used by journal editors, reviewers, funding agencies, decision-making bodies, and individual investigators.

These guidelines have been labelled as applying to clinical prognostic studies. Not all of the elements apply to studies conducted in earlier phases of marker development (Hammond and Taube, 2002), for example early marker studies seeking to correlate a new marker with other clinical variables or existing prognostic factors. However, our recommendation is that investigators conducting early marker studies should strive to adhere to as many of the reporting guidelines as applicable in their situation, and the guidelines might also suggest issues that will be important for them to consider in planning follow-up studies on their investigational markers. Studies of markers that can be used to predict the success of particular therapies, such as molecular-targeted therapies, need additional considerations. It is our opinion that predictive marker studies should generally be conducted within randomised trials, require a sufficient (usually larger) effective sample size, and assays should be in a more advanced state of development. The CONSORT statement for randomised clinical trials can serve as a starting point for reporting guidelines for predictive marker studies, but additional issues relating to the marker assays must be addressed. It is our feeling that more stringent and specific guidelines need to be developed for reporting studies of predictive markers. Such studies will be considered in somewhat more detail in the planned explanatory paper.

It may not be possible to report every detail for every study. For example, it is often difficult to provide detailed patient inclusion/exclusion criteria or treatment information in retrospective prognostic marker studies using archived tumour specimens. The impact of such missing information must be judged in the specific context of the study and its stated conclusions. For example, a 'pure' prognostic study should be conducted in a group of patients who have not received any systemic adjuvant therapy, but treatment information is often missing or unreliable in retrospective studies. In these cases, it is important to recognise that apparent 'prognostic' effects may be influenced by potential treatment by marker interactions. The key point is that there must be a clear statement of what is and what is not known. In addition, it was beyond the scope of these guidelines to recommend specific details that should be reported for each of the major classes of marker assays, for example, immunohistochemistry, *in situ* hybridisation methods, or DNA-based assays. There is an ongoing effort to define such assay-specific checklists by another working group evolving from the NCI-EORTC International Meetings on Cancer Diagnostics.

Some of the reviewers suggested that the guidelines should promote full public access to data, possibly even individual-level data. We have chosen not to include this issue in the current scope of the guidelines even though we view movement in this direction as generally positive. One concern is that if a study was

poorly designed or inadequately reported, making its data publicly available may simply propagate bad science. Good study design and data quality have to come first. We do recognise the potential benefits of promoting full public access to good quality data. It would allow verification of published analysis methods and results and would facilitate alternative analyses and meta-analyses. Attainment of these goals would be helped significantly if guidelines 10 and 11 were strictly applied, so that statistical analysis methods were described in sufficient detail to allow an individual independent of the original research team to reproduce the results of the study if supplied with the raw data. For extensive analyses, it is possible that some of this information would have to be provided as supplementary material available outside of the main published report, for example, on the journal's or author's website.

While some might view adherence to these guidelines as yet another burden in trying to publish or obtain funding, we would argue that use of these guidelines is more likely to *reduce* burdens on the research community. Making clear what is considered relevant and important to report in journal articles or funding proposals will likely reduce review time, reduce requests for revisions, and help to ensure a fair review process. Furthermore, we consider it as a prerequisite for a thoughtful presentation and interpretation of the results of a specific study and a key aid for a summary assessment of the effect of a marker in a review paper. Most importantly, what greater reduction in burden could there be than to eliminate some of the false leads generated by poorly designed, analysed, or reported studies which send researchers down unproductive paths, wasting years of time and money?

The ultimate usefulness of these guidelines will rely on how widely they are adopted. We are heartened by the enthusiastic responses we received from the several journals who have agreed to simultaneously publish this paper. There is a clear recognition in the community that the time has come (if not long overdue) to improve the quality of tumour marker study reporting and conduct. We hope that many journals will adopt these guidelines as part of their editorial requirements. To the extent that does not happen immediately, we have to rely on authors of journal articles and reviewers of those articles to initiate the movement toward adherence to these guidelines.

We expect that just as tumour marker research will evolve, these guidelines will have to evolve to address new study paradigms and new assay technologies. It is our hope that publication of these guidelines will generate vigorous discussion leading to continually improved versions and ultimately to improved quality of tumour marker studies.

The guidelines presented in this paper are available at <http://www.cancerdiagnosis.nci.nih.gov/assessment/progress/clinical.html>, as will be other recommendations from the group in due course. As noted, a detailed explanatory paper is in preparation, following the model of similar articles relating to the CONSORT and STARD statements (Altman *et al*, 2001; Bossuyt *et al*, 2003b).

## ACKNOWLEDGEMENTS

We are grateful to the US National Cancer Institute and the European Organisation for Research and Treatment of Cancer for their support of the NCI/EORTC International Meetings on Cancer Diagnostics from which the idea for these guidelines originated. We thank the UK National Translational Cancer Research Network for financial support provided to DG Altman.

## REFERENCES

- Altman DG (2001a) Systematic reviews of evaluations of prognostic variables. In *Systematic Reviews in Health Care. Meta-Analysis in Context* Egger M, Davey Smith G, Altman DG (eds) 2nd edn, pp 228–247. London: BMJ Books
- Altman DG (2001b) Systematic reviews of evaluations of prognostic variables. *BMJ* **323**: 224–228
- Altman DG, De Stavola BL, Love SB, Stepniowska KA (1995) Review of survival analyses published in cancer journals. *Br J Cancer* **72**: 511–518
- Altman DG, Lausen B, Sauerbrei W, Schumacher M (1994) Dangers of using 'optimal' cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst* **86**: 829–835
- Altman DG, Lyman GH (1998) Methodological challenges in the evaluation of prognostic factors in breast cancer. *Breast Cancer Res Treat* **52**: 289–303
- Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T, for the CONSORT Group (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* **134**: 663–694
- Bast Jr RC, Ravdin P, Hayes DF, Bates S, Fritsche Jr H, Jessup JM, Kemeny N, Locker GY, Mennel RG, Somerfield MR, for the American Society of Clinical Oncology Tumor Markers Expert Panel (2001) 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. *J Clin Oncol* **19**: 1865–1878
- Biganzoli E, Boracchi P, Marubini E (2003) Biostatistics and tumor marker studies in breast cancer: design, analysis and interpretation issues. *Int J Biol Markers* **18**: 40–48
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC (2003a) Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Standards for Reporting of Diagnostic Accuracy. *Clin Chem* **49**: 1–6
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG (2003b) Standards for reporting of diagnostic accuracy. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem* **49**: 7–18
- Brundage MD, Davies D, Mackillop WJ (2002) Prognostic factors in non-small cell lung cancer: a decade of progress. *Chest* **122**: 1037–1057
- Burke HB, Henson DE (1993) Criteria for prognostic factors and for an enhanced prognostic system. *Cancer* **72**: 3131–3135
- Burton A, Altman DG (2004) Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. *Br J Cancer* **91**: 4–8
- Concato J, Feinstein AR, Holford TR (1993) The risk of determining risk with multivariable models. *Ann Intern Med* **118**: 201–210
- Fielding LP, Fenoglio-Preiser CM, Freedman LS (1992) The future of prognostic factors in outcome prediction for patients with cancer. *Cancer* **70**: 2367–2377
- Gancberg D, Lespagnard L, Rouas G, Paesmans M, Piccart M, DiLeo A, Nogaret JM, Hertens D, Verhest A, Larsimont D (2000) Sensitivity of HER-2/neu antibodies in archival tissue samples of invasive breast carcinomas. Correlation with oncogene amplification in 160 cases. *Am J Clin Pathol* **113**: 675–682
- Gasparini G (1998) Prognostic variables in node-negative and node-positive breast cancer. *Breast Cancer Res Treat* **52**: 321–331
- Gasparini G, Pozza F, Harris AL (1993) Evaluating the potential usefulness of new prognostic and predictive indicators in node-negative breast cancer patients. *J Natl Cancer Inst* **85**: 1206–1219
- Gion M, Boracchi P, Biganzoli E, Daidone MG (1999) A guide for reviewing submitted manuscripts (and indications for the design of translational research studies on biomarkers). *Int J Biol Markers* **14**: 123–133
- Hall PA, Going JJ (1999) Predicting the future: a critical appraisal of cancer prognosis studies. *Histopathology* **35**: 489–494
- Hammond ME, Taube SE (2002) Issues and barriers to development of clinically useful tumor markers: a development pathway proposal. *Semin Oncol* **29**: 213–221
- Hayes DF, Bast RC, Desch CE, Fritsche Jr H, Kemeny NE, Jessup JM, Locker GY, Macdonald JS, Mennel RG, Norton L, Ravdin P, Taube S, Winn RJ (1996) Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. *J Natl Cancer Inst* **88**: 1456–1466
- Hilsenbeck SG, Clark GM, McGuire WL (1992) Why do so many prognostic factors fail to pan out? *Breast Cancer Res Treat* **22**: 197–206
- Hoppin JA, Tolbert PE, Taylor JA, Schroeder JC, Holly EA (2002) Potential for selection bias with tumor tissue retrieval in molecular epidemiology studies. *Ann Epidemiol* **12**: 1–6
- McGuire WL (1991) Breast cancer prognostic factors: evaluation guidelines. *J Natl Cancer Inst* **83**: 154–155
- McShane LM, Aamodt R, Cordon-Cardo C, Cote R, Faraggi D, Fradet Y, Grossman HB, Peng A, Taube SE, Waldman FM, and the National Cancer Institute Bladder Tumor Marker Network (2000) Reproducibility of p53 immunohistochemistry in bladder tumors. *Clin Cancer Res* **6**: 1854–1864
- McShane LM, Simon R (2001) Statistical methods for the analysis of prognostic factor studies. In *Prognostic Factors in Cancer* Gospodarowicz MK, Henson DE, Hutter RVP, O'Sullivan B, Sobin LH, Wittekind Ch (eds) 2nd edn, pp 37–48. New York: Wiley-Liss
- Mirza AN, Mirza NQ, Vlastos G, Singletary SE (2002) Prognostic factors in node-negative breast cancer: a review of studies with sample size more than 200 and follow-up more than 5 years. *Ann Surg* **235**: 10–26
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup D, for the QUOROM Group (1999) Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet* **354**: 1896–1900
- Moher D, Schulz KF, Altman D, for the CONSORT Group (2001) The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* **285**: 1987–1991
- Paik S, Bryant J, Tan-Chiu E, Romond E, Hiller W, Park K, Brown A, Yothers G, Anderson S, Smith R, Wickerham DL, Wolmark N (2002) Real-world performance of HER2 testing – National Surgical Adjuvant Breast and Bowel Project Experience. *J Natl Cancer Inst* **94**: 852–854
- Popat S, Matakidou A, Houlston RS (2004) Thymidylate synthase expression and prognosis in colorectal cancer: a systematic review and meta-analysis. *J Clin Oncol* **22**: 529–536
- Riley RD, Abrams KR, Sutton AJ, Lambert PC, Jones DR, Heney D, Burchill SA (2003a) Reporting of prognostic markers: current problems and development of guidelines for evidence-based practice in the future. *Br J Cancer* **88**: 1191–1198
- Riley RD, Burchill SA, Abrams KR, Heney D, Sutton AJ, Jones DR, Lambert PC, Young B, Wailoo AJ, Lewis IJ (2003b) A systematic review of molecular and biological markers in tumours of the Ewing's sarcoma family. *Eur J Cancer* **39**: 19–30
- Riley RD, Heney D, Jones DR, Sutton AJ, Lambert PC, Abrams KR, Young B, Wailoo AJ, Burchill SA (2004) A systematic review of molecular and biological tumor markers in neuroblastoma. *Clin Cancer Res* **10**: 4–12
- Roche PC, Suman VJ, Jenkins RB, Davidson NE, Martino S, Kaufman PA, Addo FK, Murphy B, Ingle JN, Perez EA (2002) Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. *J Natl Cancer Inst* **94**: 855–857
- Schilsky RL, Taube SE (2002) Introduction: tumor markers as clinical cancer tests – are we there yet? *Semin Oncol* **29**: 211–212
- Schumacher M, Holländer N, Schwarzer G, Sauerbrei W (2005) Prognostic factor studies. In *Handbook of Statistics in Clinical Oncology* Crowley J (ed) pp 307–351. New York: CRC Press (Chapter 18)
- Simon R (2001) Evaluating prognostic factor studies. In *Prognostic Factors in Cancer* Gospodarowicz MK, Henson DE, Hutter RVP, O'Sullivan B, Sobin LH, Wittekind Ch (eds) 2nd edn, pp 49–56. New York: Wiley-Liss
- Simon R, Altman DG (1994) Statistical aspects of prognostic factor studies in oncology. *Br J Cancer* **69**: 979–985
- Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB (2000) Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* **283**: 2008–2012
- Thor AD, Liu S, Moore II DH, Edgerton SM (1999) Comparison of mitotic index, *in vitro* bromodeoxyuridine labeling, and MIB-1 assays to quantitate proliferation in breast cancer. *J Clin Oncol* **17**: 470–477